Relating Planner Task Performance for Container Terminal Operations to Multi-Tasking Skills and Personality Type

Shalini Kurapati[a*] , Heide Lukosch[a], Stephanie Eckerd[bc], Alexander Verbraeck[ab], Thomas Corsi[b]

[a] Delft University of Technology, Jaffalaan 5, 2628 BX, Delft, The Netherlands

[b] University of Maryland, 3359 Van Munching Hall, College Park, MD, 20742 United States

[c] University of Tennessee, 310 Stokely Management Center, Knoxville, TN, 37996-0530 United States

*corresponding author.

Corresponding mail address: Jaffalaan 5, 2628 BX, Delft, The Netherlands

Corresponding e-mail address: S.Kurapati@tudelft.nl

Abstract

Planning the operations within a container terminal is a complex task. It requires planners to demonstrate adaptive behavior while handling stressful, complex, and unexpected situations in today's dynamic and technology dependent workplace. This paper aims at investigating the role of multi-tasking ability, moderated by an individual's personality type, in predicting planner task performance using simulation gaming methods. Hierarchical regression analysis results demonstrate that the direct effect of multi-tasking ability on performance is positive and significant. With one exception, the personality traits do not significantly intensify or lessen the impact of multi-tasking in predicting task performance. The personality trait, openness to experience, significantly lessens the impact of multi-tasking ability on performance. Our results suggest that container terminal operators may benefit by considering the above-mentioned results while allocating planning tasks to their employees and new recruits. The instruments used in this research could also be used for evaluating and training candidate planners.

Keywords: Planner task performance, container terminals, job performance, multi-tasking ability, personality type, simulation gaming

1. Introduction

Planning of logistic activities in seaports has undergone a paradigm shift since the emergence of container terminals in the 1960s. As container terminals increased, its logistics planning tasks have become more complex, dynamic, inter-dependent, and technology dependent (Murty, Liu, Wan, & Linn, 2005). Logistic activities within container terminals are costly as they require the combined use of expensive resources such as cranes, terminal equipment, and manpower. The competitiveness of a container terminal is determined by its service level to its clients (e.g. shipping lines) and the efficient use of its resources (Legato & Mazza, 2001; De Oliveira & Cariou, 2015). However, the smooth and cost-effective flow of containers through a container terminal requires an integrated planning approach, which is systematic and comprehensive in nature (Meisel, 2009).

Employee performance has always been the backbone of container terminal productivity (Legato & Monaco, 2004). Terminal employees mainly consist of 'ground' workers, also called 'core' workers (Legato & Monaco, 2004; Turnbull & Weston, 1993). Ground workers are usually involved with loading and unloading operations as well as housekeeping activities in the terminal. They are each responsible for a single main task, depending on their experience and skills. Typical ground workers are crane operators, checkers, deck men, straddle-carrier drivers, terminal truck drivers, and raisers (Legato & Monaco, 2004). However, with the advent of modern and automated cargo handling systems, the ground workforce may have to move away from single tasks, creating a dire need for a renewed workforce, well-trained and skilled to run the terminal (Turnbull & Weston, 1993). However, automated handling systems and related rapidly progressing port technologies put a lot of pressure on terminal operators and employees as the skill-set required for superior performance has not been well explored yet (Notteboom, 2012).

Few studies have paid attention to the skills associated with superior performance in *planning tasks* of modern terminals. Studies done by Gordon, Lee, and Lucas (2005) mention the positive effect of general skills, such as operational skills and IT management skills on planner task performance. Notteboom (2012) identifies the need for adaptive planning and a multi-functional workforce being able to multi-task to deal with the complex work environment of the modern container terminal. In other fields, such a dynamic, complex, and technology dependent work environment requires employees with adaptive skills (Burke, Pierce, & Salas, 2006). Adaptive performance is characterized by the ability to handle dynamic situations, to deal with stressful events, to manage crisis situations, and to navigate unfamiliar or unpredictable work situations (Pulakos, Arad, Donovan, & Plamondon, 2000). The challenge for container terminal organizations is to identify and train employees who can achieve highest performance in the dynamic, challenging work environment, since the performance of employees influences the productivity and competitiveness of organizations.

Hence, it would be significant if those employee attributes can be identified that relate to superior planner task performance. We see this as the core knowledge gap addressed in this paper. As noted above, given the need for adaptive skills, it would be relevant to explore the link between the ability to simultaneously address challenges, or the ability to multi-task, and success in planner task performance. Furthermore, it would also be relevant to explore general personality characteristics and their links to planner task performance as they have been identified in literature as predictors for adaptive performance in other fields (Penney et al., 2005). This represents an important and as yet underexamined interface between the logistics and human resources disciplines (Swart, Hall, & Chen, 2012).

This research presents the results of a series of experiments that address the knowledge gap. In this research, we seek to identify key behavioral attributes relevant to planner task performance, thereby contributing to a stream of research addressing the interaction between logistics and HR disciplines (McAfee, Glassman, & Honeycut, 2002; Lu & Kuo, 2016). Our experiments include tools that test the skills of individuals to efficiently multi-task; to support planning task ability as well as to identify individual personality characteristics. The research presents a combination of the Multi-Attribute Test Battery (MATB-II) (Santiago-Espada, Myer, Latorella, & Comstock, 2011) to measure multi-tasking abilities, a simulation game called Yard Crane Scheduler (Kurapati, Groen, Lukosch, & Verbraeck,

2014) to measure planner task performance for container terminals and a Big-Five model personality identification questionnaire (Gosling, Rentfrow, & Swann, 2003). The aim of this combined approach is to gather deeper insights in the role of multi-tasking skills and personality traits when it comes to the execution of planning tasks in a complex, interdependent, and dynamic environment. The remainder of this article is structured as follows. First, the complexity of the planning tasks in modern container terminals is explained in more detail. Then, the theoretical background of multi-tasking skills and personality types is discussed and two research hypotheses are derived based on the knowledge gap. Following the research hypotheses, the methodology section explains the experimental set-up and the various research instruments used for the measurements. The subsequent section is dedicated to the presentation and analysis of results. The final section draws conclusions and recommendations along with providing a discussion related to the limitations and the future of the current study.

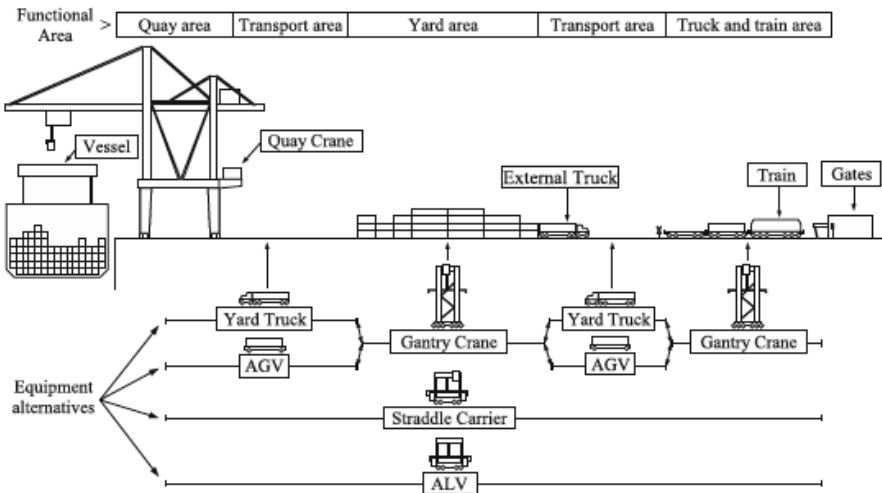## 2. Complexity of Planning Operations in Container Terminals

This section provides some background on container terminal operations as well as the need to identify employee skills and traits for superior planner task performance.

2.1. Planning operations in container terminals

Container trade rose by 600% in the last 20 years (UNCTAD, 2014). Over 80% of the volume of the global trade is seaborne and 70% of the value of dry cargo is transported using containers (UNCTAD, 2014). The mismatch between ever-increasing volumes of containers arriving in container terminals and relatively inert port infrastructure and size make the planning operations in container terminals very complex (Meisel, 2009). A schematic overview of a typical container terminal is presented in Figure 1.

Container ships or vessels carrying containerized goods arrive at a container terminal to offload export and transshipment containers and to load import containers. Quay cranes serve these ships. The containers that are unloaded are carried to the yard, where the containers wait to be carried into the hinterland with their next mode of transportation, which could be truck, train, or barge. The reverse process applies to the containers that are loaded on to the ship (Gambardella, Rizzoli, & Zaffalon, 1998). Containers are transported within the terminal by trucks, tractor trailers, straddle carriers, or Automated Guided Vehicles (AGVs). The following sub-section explains the planning tasks in a container terminal.

Figure 1: Schematic processes in a container terminal together with equipment (Meisel, 2009)

2.1.1. Complexity in container terminal operations and need for adaptive planning

The planning tasks in (modern) container terminals are clearly illustrated in Figure 2. As one can observe all the planning tasks related to seaside, yard, and land side operations are inter-dependent, meaning that a change in one of the plans can have major effects on other plans, possibly leading to a cascading effect.

Figure 2: Planning tasks in container terminals (Meisel, 2009)
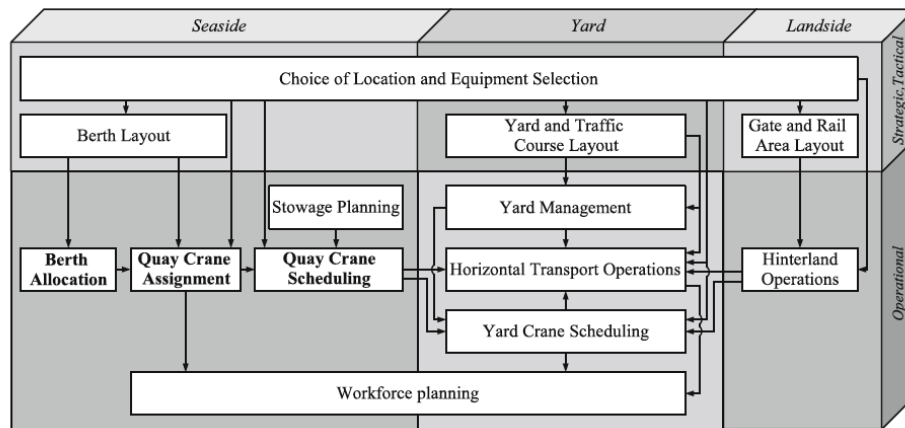
The high volumes of containers make these planning tasks very complex (Meisel, 2009). Planning in container terminals takes place under dynamic circumstances. The frequent changes in the schedules of ships, changes in truck arrival times due to traffic, customs clearance, and other disruptions such as breakdowns of equipment and bad weather conditions contribute to the dynamics. This complexity and dynamics call for adaptive and flexible planning in container terminals. However, in practice each of the planning tasks is decomposed into sub planning tasks and optimized piece by piece and often sequentially. The literature overview of Carlo, Vis and Roodbergen (2014) shows that even in literature, different planning operations are addressed on a one-by-one basis. One of the research avenues they propose is to "integrate storage space assignment and QC scheduling" (p.428). In the sequential solution process, decisions made are static and cannot be changed easily, and several ad-hoc modifications are needed when a problem is identified or when poor performance is detected. These modifications disturb the operations of container terminals by increasing berthing and vessel handling times and reduce the reliability of a terminal. This negatively affects customer satisfaction, and thereby the competitive advantage of the terminal (Meisel, 2009; De Oliveira & Cariou, 2015).

The described complexity of container terminal operations calls for planning to be done in an adaptive and integrated manner to deal with dynamic circumstances and to secure the market position of the terminal (Meisel, 2009). Therefore, the focus shifts to planners who are responsible for and able to carry out this complex task.

In general, distinct factors influencing performance of individuals in complex task settings have been described in literature already. Such factors comprise a diverse range of individual traits, with variables mainly spanning six major categories: personality, demographics, physiology, cognition, motivation and skills (Baines et al., 2005). Task performance reflects activities that are formally recognized as part of the job, support the organization's technical core, and directly impact organizational goal accomplishment (Morgeson, Reider, & Campion, 2005). Cognitive ability has a positive effect on job performance in general (LePine & Van Dyne, 2001). To perform in complex systems, people always have to continually monitor situations and adjust behavior according to changing task demands (Rouse, Cannon-Bowers, & Salas, 1992). Although it is known that talented and skillful workforce contribute to the performance and efficiency of container terminals, the role of personnel selection and human resources management in container terminal performance is still

considered as non-traditional as they are difficult to quantify (Yoon, Lee, & Dinwoodie, 2015). Although a small number of studies (e.g., Notteboom, 2012) provide minimal insights into the general skills supporting complex task performance such as IT management skills, to the best of our knowledge the relationship between individual skills, personality traits, and planner task performance is largely unknown. In the following section, we explore the role of multi-tasking ability and personality type in relation to planner task performance as described more or less in isolation in literature thus far. The ensuing hypotheses based on this problem description gap analysis are presented together with further theoretical background on each of the two relationships in the following sections.

3. Multitasking skills and task performance

The increasing adoption of Information and Communication Technologies (ICT) since the 1990s has made multi-tasking indispensable at the workplace (Appelbaum, Marchionni, & Fernandez, 2008). Computer-based multi-tasking skills in particular have become a prerequisite for most job descriptions. Companies find multi-tasking a highly desirable skill in prospective employees (Adler & Benbunan-Fich, 2012; Appelbaum et al., 2008). As most of the jobs today require at least some degree of multi-tasking ability, good performance is often associated with good multi-tasking ability at the workplace (Bühner, König, Pick, & Krumm, 2006). However, there are several definitions and interpretations of multi-tasking in literature. As all planning tasks in a container terminal are supported by ICT, we focus our interest on computer-based multi-tasking.

In our research context, multi-tasking is defined as the ability of an individual to perform several independent yet concurrent computer-supported tasks (Adler & Benbunan-Fich, 2012). The degree of concurrence varies between sequential, parallel, and interleaved type of tasks. Sequential tasks have almost zero concurrency, but they set the base line for multi-tasking. Parallel tasks represent true multi-tasking, but humans are not able to fully divide their attention to perform different tasks in parallel (Adler & Benbunan-Fich, 2012). This kind of parallel-tasking is also called dual-tasking, when people attempt to perform two different tasks at the same time (Pashler, 2000). Related to this is the relatively new concept of 'continuous partial attention', introduced by (Stone, 2007). The difference between both concepts lies, according to Stone (2007), in the motivation. While people perform multi-tasking out of the desire to be more productive and/or efficient, they pay partial attention to simultaneous tasks out of the wish to be connected, and not to miss out any opportunities or information (Rose, 2010). Another definition of this concept is discussed by (Foehr, 2006), who describes the phenomenon of continuous partial attention as "media multi-tasking" (p.1), thus describing a situation where one engages in more than one medium, e.g. reading onscreen on a computer while communicating via a cell phone. Interleaved tasks best describe human multi-tasking, where attention to one task is temporarily suspended to allocate attention to the other tasks. In this way the originally task is resumed, after briefly being abandoned, after the user attends another task (Adler & Benbunan-Fich, 2012; Salvucci & Taatgen, 2010). This is supported by Oberauer and Kiegl (2004), who state that human cognitive processes do not allow for parallel cognitive operations, and by Ophir, Nass, and Wagner (2009), considering simultaneously performing multiple tasks a challenge for human cognition. In our research, we consider multi-tasking to be interleaved.

Although multi-tasking is a highly valued skill in organizations, its effect on performance has not been established thoroughly. Studies on this topic are limited and the available evidence is inconsistent and contradictory (Adler & Benbunan-Fich, 2012; Appelbaum et al., 2008). Given the importance and prevalence of multi-tasking at today's work place, we presume that multi-tasking ability can be conducive to task performance. In this study, we aim at understanding this effect specifically on planner task performance for container terminal operations. This leads us to our first hypothesis H1.

H1: Multi-tasking ability is positively related to planner task performance for container terminal operations.

In addition to multi-tasking, another individual trait that predicts performance is personality type. The following section provides the theory and motivation to investigate the role of personality type in planner task performance.

4. Role of personality type in task performance

Personality can be defined as "the sum of physical, mental, emotional and social characteristics possessed by a person that uniquely influences his cognitions, motivations, and performance in any environment" (Othman, Bhuiyan, & Gouw, 2012). Personality type is seen as a very important behavior moderator of employee performance in organizations (Pew & Mavor, 1998). Since its early conceptualization, there is an agreement that for personality, five factors can be distinguished. Empirical studies have shown that most researchers (and 'common' people, too) use similar terms to describe personality (Digman, 1990). Especially in the 1990s, the work of Digman (1989) and Goldberg (1990) consolidated the use of a five-factor model to describe personality, also called the "Big Five" model. Direct effects of personality have been studied since the 1960s, although the prediction of job performance using personality type has been long debated (Penney et al., 2011). However, a multitude of empirical studies strongly supports the predictive capability of the Big Five model of personality with respect to job performance. If applied with due diligence, the utility of the Big Five model to predict job performance of employees is seen as very valuable (Penney et al., 2011; Hurtz & Donovan, 2000). This led to the renewed interest among researchers to explain the direct and indirect effects of personality type on performance. Regardless of the debate, personality type remains a very important determinant of employee behavior and job performance in the workplace. Although personality has been conceptualized from various theoretical perspectives, researchers largely agree on the Big Five model of personality. The Big Five traits don't represent a single theoretical perspective but have been derived from the analyses of natural language terms that people use to describe themselves (Pervin & John, 1999). The Big Five model of personality constitutes the degree to which every individual varies within the following five traits (Goldberg, 1990).

(1) Extraversion: An individual scoring high on extraversion would typically be quite talkative, assertive, sociable, adventurous, energetic etc.;

(2) Agreeableness: An agreeable person can best described as good natured, generous, tolerant, honest, courteous, etc.;

(3) Conscientiousness: Individuals with high conscientiousness are usually orderly, consistent, responsible, mature, sophisticated, formal, reliable, dependable, etc.;

(4) Emotional Stability: Individuals with high emotional stability are typically calm, poised, self-reliant, etc.;

(5) Openness to Experience: Individuals who are more open to new experiences are usually creative, original, curious, artistic, imaginative, independent, etc.

The Big Five model has been widely used for employee selection and assessment by human resource professionals. However, it is important to define the criterion for job performance to establish empirical validity (Penney et al., 2011). The three broad categories used for assessing job performance are task performance, contextual/citizen performance, and counter-productive behavior (Rotundo & Sackett, 2002). Task performance encompasses all the formally recognized activities related to a certain job that contribute to the organization's technical core, in other words, all job related activities that lead to increased production of goods or provision of services. The contextual or citizen performance relates to the interpersonal behavior within teams or at social situations at work. Counterproductive behavior consists of all activities that compromise the integrity of the organization (Rotundo & Sackett, 2002). In this study we focus on the job performance related to task performance alone. Task performance in complex systems can be characterized by the following activities (Pew & Mavor, 1998):

- Assessing the current situation,
- Activating new tasks in response to recent events,
- Assessing task status to verify whether each task is being performed well,
- Terminating tasks with achieved or unachievable goals,
- Assessing task resource requirements,
- Prioritizing active tasks,
- Allocating resources to tasks in order of priority, and
- Updating the task agenda.

The set of activities above has been adapted from cockpit task management and is generalized to other adaptive task performance in complex systems (Pew & Mavor, 1998). It can be observed from the above activities that task performance requires task switching, prioritizing, and resource allocation, which are very much related to multi-tasking ability. Therefore we are very interested to understand the moderating effects of personality on the relationship between multi-tasking ability and task performance. A moderating variable is a quantitative or a qualitative variable that affects the strength and direction of the relationship between two variables (Baron & Kenny, 1986). As discussed earlier, although the direct effects of personality on job performance have been debated (Penney et al., 2011), researchers agree that personality type "influences" the performance of individuals in a given environment (Penney et al., 2011; Hurtz & Donovan, 2000; Othman, Bhuiyan, & Gouw, 2012). Therefore, we choose to consider the moderating effects of personality traits rather than its direct effects for our research. This brings us to our second hypothesis,

H2: Personality traits, a) Extraversion b) Agreeableness c) Conscientiousness d) Emotional stability e) Openness to experience, have a moderating effect on the relationship between multitasking ability and planner task performance for container terminal operations.

5. Research design and methodology

Our research methodology includes the setting of the research instruments used in our study, the selection of participants, the experimental set-up and sampling procedure, and the analysis of the data. The data analysis is presented in the results section.

5.1 Sample

Our experiments have been conducted in a controlled setting, with classes of 37 bachelor's and 96 master's degree students from the transport and logistics domain, from universities in the Netherlands and the United States. The nationalities of the students included Dutch, Chinese, American, Indian, Finnish, German, Greek, Columbian, Costa Rican, Taiwanese, South Korean, Syrian, Vietnamese, Pakistani and Ecuadorian. We assume that these students will be future professionals in the supply chain, logistics, and transportation field and believe that they will encounter similar complex planning issues like container terminal planning tasks in their future jobs. Students fall within the theoretical boundary conditions of what is being tested, and so we utilize them for this research (Thomas, 2011). This supports our research study that considers students to be a proxy for professionals in the field of supply chain, logistics and transportation. Sixty-eight students were male, and 65 students were female. Although a total of 142 students participated in the actual study, we could use only the data from 133 students, due to incomplete surveys and student refusal to provide data to the study. We received the approval of the Institutional Review Boards of the involved American and Dutch Universities to conduct this research study with students.

5.2 Procedure

Participation in the experiment was voluntary, and the participants were incentivized to perform well by the possibility to win a small prize through a weighted lottery method based on their task performance score. Each experimental session took about 2 hours and 15 minutes to execute. Participants were asked to complete a pre-survey that collected their demographic information as well as their personality traits using the Ten-Item Personality Inventory (TIPI) test, which measures the Big Five dimensions (Gosling et al., 2003). Following the pre-survey, the participants were provided with a detailed briefing for their multi-tasking exercise and they were asked to execute the exercise using a software tool known as the "Multi Attribute Test Battery" (MATB-II) (Santiago-Espada et al. 2011). After MATB-II, the participants were presented with a brief lecture on container terminal planning operations, and they received a tutorial and practice session for the "Yard Crane Scheduler" (YCS) simulation tool that was used to assess their planning task performance (Kurapati et al., 2014). Two exercises with varying levels of difficulty (Mission 1 and Mission 2) were provided to the participants after the tutorial session. The participants were asked to complete mission 1 twice to get acquainted with the YCS environment and the controls. They were allowed to ask questions and clear any questions regarding the functionality of YCS. The scores of Mission 1 were not counted for the data analysis. The participants then had to complete the more difficult Mission 2 twice. The highest score of Mission 2 was considered for evaluating the planning task performance of each participant. After the YCS exercise the facilitator conducted a de-briefing session to gather their insights and strategies related to YCS and to discuss the challenges in container terminal planning operations.

5.3. Instruments and Measures

For the experiments, we provided the needed hardware and software to the students. All experiments were held within a classroom setting. Each session was managed by one teacher and at least two researchers providing observation and technical support. As we introduced above, the experiments were meant to explore the relationship of personality traits, multi-tasking ability, and complex planner task performance. For each of the variables, we used a distinct measurement method, which we introduce in the following.

5.3.1. Measuring personality traits: the TIPI questionnaire

Several extensive tests are available for the measurement of personality traits. Examples include the 240-item NEO Personality, Inventory, Revised (NEO-PI-R), the 44-item Big Five Inventory test (BFI), and the 100-item Trait Descriptive Adjective test (TDA) (Gosling et al., 2003). Most of these tests are time consuming, expensive and may sometimes need a trained professional to administer them. The design and set-up of the research motivated us to seek an openly available personality test that is very short but well tested and validated. Our search led us to the Ten-Item Personality Inventory (TIPI). As with every brief measure, somewhat diminished psychometric properties are the main limitations of TIPI (Gosling et al., 2003). However, the test has widely been adopted by the research community with over 1,000 citations since its publication. Participants are asked to describe their personality with 10 adjectives pairs. They have to choose a number between 1 and 7 for each pair, with 1 signifying total disagreement and 7 meaning total agreement. The scores for each of the Big Five traits are calculated based on the responses.

5.3.2. Measuring multi-tasking ability: MATB-II

MATB-II task descriptions

The Multi Attribute Test Battery (MATB-II) was chosen as the research instrument for measuring multi-tasking ability. MATB was developed in the early 90's for laboratory studies of operator performance and workload, including tasks representational for the aviation domain (Comstock & Arnegard, 1992). MATB was upgraded to MATB-II following a wide acceptance from research communities not limited to space and aircraft applications. MATB and MATB-II have been used in over 130 research studies on human performance related to multi-tasking and complex tasks (Santiago-Espada et al., 2011). Based on our search for a validated tool to capture multi-tasking ability, we found that MATB-II is an apt tool that clearly represents computer-based multi-tasking of complex tasks. The software was requested from the subsidiary website of NASA (http://matb.larc.nasa.gov/). As illustrated in Figure 3, MATB-II is a computer based multi-tasking exercise, which requires the simultaneous performance of systems monitoring, tracking, dynamic resource management, and communication tasks.

Figure 3: Screenshot of the various tasks in the MATB-II exercise (Santiago Espada et al., 2011)
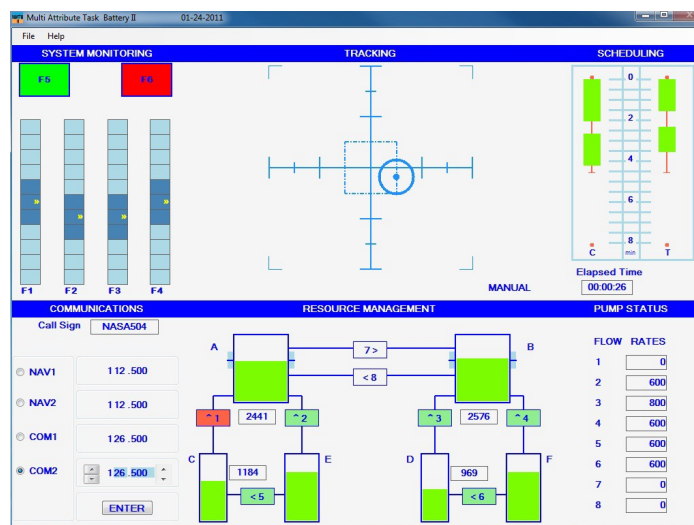
The following four tasks of MATB-II were used in the study:

SYSMON task. The systems monitoring task or SYSMON is in the upper left corner of the MATB-II task window. SYSMON has two sub-tasks: detecting changes in the warning lights and changes in the scales of the system, which ask the participants to clear warning lights and ensure that values for the scales stay within the boundaries and respond if they do not. In our study, we offered the choice of using either the mouse or the keyboard to control this task. The performance of the SYSMON task is determined by the (lowest) number of errors each participant commits in terms of timely noticing changes in the boxes and scales and reacting correctly to them.

TRACK task. The tracking task or TRACK is in the upper central part of the MATB-II task window. This tasks has two modes: MANUAL or AUTO ON. If the MANUAL mode is on, the participants need to keep the target in the center of the inner box using a joystick. By default, the software records the position of the target with respect to the center point every 15 seconds, by calculating the Root Mean Square Distance from the Center (RMSDC) of the target from the center point. The closer the participant positions the target to the center, the better they perform at the TRACK task.

RESMAN task. Dynamic resource management or RESMAN task is in the lower center part of the MATB-II task window. RESMAN represents a fuel management system containing five tanks and a set of 8 pumps. The objective of the RESMAN task is to maintain the fuel levels in Tank A and Tank B as close to 2500 units of fuel as possible throughout the exercise, while fuel demand is changing. To maintain this objective, participants have to switch pumps on and off, with pumps breaking down once in a while. Each pump has different flow rates.

9

COMM task. The communications or COMM task can be found in the lower left corner of the MATB-II task window. Participants are provided with ear buds through which they can hear auditory commands generated by the MATB-II software that requests them to choose a certain radio and change its frequency. Participants are only expected to respond to the auditory commands that have a given call sign. Any other command has to be ignored.

The other aspects of the MATB-II task window include a scheduling display, SCHED, that indicates the communication and tracking tasks' timeline during the exercise. This display does not require participant intervention. During the exercise, a workload rating scale based on NASA-TLX is displayed while the run timer of the exercise freezes. After the participants fill in the rating scale, the tasks resume with the previously saved states of each of the tasks.

Each of the four tasks of MATB-II has a different output file with varying performance indicators. The following section describes our efforts to define a composite score for MATB-II.

MATB composite score

Although many studies have used MATB and MATB-II to analyze multitasking abilities and workload, most of their findings were based on the individual performance indicators of each of the four tasks in the MATB-II exercise (Harris, Hancock, Arthur & Caird, 1995; Fairclough, Venables & Tattersall, 2005; Caldwell & Ramspott, 1998). We decided to design a composite MATB score that is more valuable at measuring the combined performance of all the four tasks instead of measuring the performance for individual tasks which is counterintuitive to the concept of multi-tasking. We modeled the composite MATB score as the sum of the four task performance scores. We made sure that the task performance scores were normalized on a scale from zero to one, resulting in an overall score of zero to four. A high overall score can only be obtained when participants do well on all four tasks by paying attention to the status of the different tasks at the same time and simultaneously reacting on the visual and auditive cues that are provided. Participants had to use motor skills (joystick, keyboard, mouse), recognition skills (colors, gauges, call sign), tracking skills (follow target, view fuel level changes, view pump statuses), and calculation skills (which pumps to open and close) in parallel. The task performance scores were calculated by measuring the performance of each of the tasks in terms of deviations from task goals. The deviations in SYSMON and COMM are quite straightforward to understand – the task goal is no errors, so the number of errors accounts for the deviation – but we had to clearly define them for RESMAN and TRACK consistent with their task goals. Each of the task performance scores is calculated as follows.

(1) **SYSMON score**

MATB measures the variables S_T, the number of timeouts, or lack of response, during the systems monitoring task, and S_E, the number of errors related to the response to changes in the boxes and scales. Participants are given a two second grace period to avoid counting double clicks as two errors. The total number of errors $S = S_T + S_E$ is normalized to S_{norm} where $0 \leq S_{norm} \leq 1$, by dividing it by the highest number of possible errors, S_{MAX}, defined as the number of errors in the 'do nothing' scenario. This score is obtained when participants either do not react to any of the system monitoring tasks, or choose the wrong reaction on every one of them. If necessary, the normalized score is capped between zero and one as participants can theoretically make more errors than the number of tasks. In that case they will receive a zero score. S_{norm} is now defined as (higher is better):

$$S_{norm} = 1 - \min\left(1, \frac{S_T + S_E}{S_{MAX}}\right)$$

(2) **TRACK score**

MATB measures the deviation T_i from the center point every 15 seconds, or 20 times in total during the five-minute experiment. The root-mean-square deviation T_{RMS}, is calculated and normalized to T_{norm}, where $0 \leq T_{norm} \leq 1$. If T_{norm} is one, then the target was perfectly positioned on the center point during the entire exercise. The maximum outlier score T_{MAXRMS} is the root-mean-square of

the distance of the cursor position from the center point, when the participants completely ignore the tracking task ('do nothing' scenario). The tracking score is modeled as (higher is better):

$$T_{norm} = 1 - \min\left[1, \frac{\sqrt{\frac{1}{20}\sum_{i=1}^{20}(T_i)^2}}{T_{MAXRMS}}\right]$$

(3) RESMAN score

MATB measures the deviation from the required fuel levels of 2,500 liters in both tanks A and B every 30 seconds during the exercise, or 10 times in total. Assume that $R_{A,i}$ indicates the fuel level for tank A at sample i, and $R_{B,i}$ indicates the same for tank B. We can then calculate the absolute fractional deviation $F_{T,i}$ of each tank content compared to the 2,500 liter target level. The normalized deviation score R_{norm} (higher is better) is calculated by dividing the sum of $F_{A,i}$ and $F_{B,i}$ over 10 observations by R_{MAX}, the deviation when no action is taken by the participant, and capped at zero in case the participant does worse than the 'do nothing' scenario.

$$R_{norm} = 1 - \min\left[1, \frac{\sum_{i=1}^{10}\left(\left|\frac{R_{A,i} - 2500}{2500}\right| + \left|\frac{R_{B,i} - 2500}{2500}\right|\right)}{R_{MAX}}\right]$$

(4) COMM score

MATB measures the number of communication errors C, e.g. as a result of time outs, wrong radio choice, wrong frequency choice or a response to the wrong call sign. C is normalized to C_{norm}, where $0 \leq C_{norm} \leq 1$ by dividing it by the number of communication tasks C_T to which the participant should react, plus the number of 'false' call signs C_F to which the participant should not react. The COMM score, where higher is better, is then defined as:

$$C_{norm} = 1 - \min\left[1, \frac{C}{C_T + C_F}\right]$$

(5) MATB_COMP

The composite score of multi-tasking ability MATB_COMP score is defined on a scale from 0 to 4 as the sum of the standardized individual task scores:

MATB_COMP = SYSMON score+ TRACK score+ RESMAN score+ COMM score

$$= S_{norm} + T_{norm} + R_{norm} + C_{norm}$$

5.3.3. Measuring planner task performance for container terminal operations: the YCS game

To create a controllable experimental set-up for the measurement of planner task performance, we decided for a simulated container terminal environment, namely a simulation game called Yard Crane Scheduler (YCS). We see a strong potential for simulation games as research instruments since games produce an abundance of data related to user experience, decision-making, human behavior, and skills. Scenarios in games provide hands-on experiences with real-life problem solving skills of the participants (Niehaus & Riedl, 2009) in safe experimental environments. Games can be used in a controlled fashion with a large group of users, compared to a small number of observations in an uncontrolled environment for real planers in container terminals.

When games are used for research, the validity, or degree of correspondence, between the reference system and the simulated model thereof is crucial (Raser, 1969; Peter, Vissers, & Heijne, 1998). Especially computer-based games offer the possibility to retrieve data on player behavior in order to analyze them. The player behavior leads to a certain score within the game that can be used to measure performance (Harteveld & Sutherland, 2015). This is the case when scores are related to certain learning objectives of the game, allowing a formative assessment of performance. In the case of the game we used to study performance of planning tasks, experts from the field of container transportation validated this relationship (Kurapati et al., 2014). Moreover, the scores of the game were also used as informational feedback to motivate the participants (Harteveld & Sutherland, 2015).

Background and objective of the YCS game

Yard Crane Scheduler or YCS is a web-based game developed by researchers and a professional game developer. The development process has been conducted in close collaboration with practitioners in the container terminal industry (Kurapati et al., 2014). The main objective of YCS is to manage and align various planning operations in the container terminal in order to maximize terminal profit, indicated as an overall score. YCS focuses on the two main challenges observed in container terminal operations; dynamic planning, distribution of containers in the yard, and allocation of resources to ensure maximum utilization of terminal resources. YCS is a MicroGame, a simulation game that transfers the meaning of a real world environment in a simplified yet sufficiently realistic manner (Kurapati et al., 2014). The advantages of such MicroGames is their strong relation to real-world problems and the short timeframe in which they can be administered. That makes them suitable to be used as an experimental instrument. Naturally, any simplified version of reality has limitations with respect to applying it in practice. However, in the context of our research objective, we worked towards minimizing any such limitations by underpinning the design of YCS by a framework known as Triadic Game Design, the concept of balancing a game between reality, meaning, and playability (Harteveld, 2011). The YCS game has been extensively tested besides our controlled experiments. Student participants have played the game more than 12,700 times, while professionals in the container terminal industry played over 15,000 times so far. An extensive face validation of the game including the scoring system was done by a group of five experts in the container terminal industry with between three and 25 years of experience in marine container terminal operations.

General description of YCS

The YCS game is a web browser-based online game that is hosted on an online portal. The main screen of the game shows a simplified top view of a container terminal as shown in Figures 4 and 5.

Deep sea vessels arrive to load and unload containers in the terminal. All the containers to be unloaded from the ships are known as import containers. Export containers in the yard that need to be loaded onto the ship are marked with an arrow symbol. In this game, the yard is one tier high and loading and unloading of ships is not done simultaneously. The objective for the player is to make sure the ship is serviced as soon as possible, while making efficient use of terminal resources.

The game has three tutorials, with information required to make participants familiar with the rules and game mechanics. The game has various missions. Each mission represents the difficulty level, Mission 1 being the easiest while the complexity increases in the subsequent missions. At the end of every mission players will be shown a detailed scoring screen containing the score and details of their performance.

Resources in the game

Various types of equipment as used in container terminals are incorporated in the game, consistent with the overview of container terminal operations provided in Figure 1.

- *Quay Cranes (QCs)* are located on the quay- or waterside of the terminal. QCs load and unload containers on and off the ships. These cranes are mounted on rails and cannot pass over each other, both in reality and in the game. There are 3 QCs in the game.
- *Automated Guided Vehicles (AGVs)* are driverless trucks that transport containers from the ship to the yard and vice-versa. AGVs are automated and do not require any planning. Whenever a crane is at work, an AGV arrives at that spot to either pick up or drop off a container. There are 10 AGVs in the game.
- *Rail Mounted Gantry cranes (RMGs)* or yard cranes are located in the yard to stack containers in the yard. They pick up the containers from AGVs and place them in the yard. They also retrieve containers out of the yard and place them on AGVs, which then carry them towards the quayside. RMGs can only move along the container block as they are rail mounted. If RMGs need to be changed to a parallel block, players need to drag them until the end of a block and drag them sideways to the required parallel block.
- *Trucks* are involved with landside or hinterland transportation. In the game trucks arrive to pick up yellow containers that are designated to truck transportation.

Tasks within the game

The containers to be unloaded by every ship need planning and resource allocation on the yard side, quay side, and land side of the terminal, consistent with the planning tasks in Figure 2. The arrival times and departure times are indicated for each ship on the right hand side of the screen. There are several planning and resource allocation tasks in the game. The game functions in two modes: the planning mode and the operational mode, shown in Figures 4 and 5. For the planning mode, players need to plan the location of the import containers in the yard for all the arriving ships. The arrival times are clearly indicated on the right hand side of the screen. The color of the ship is matched with the containers that it carries. Players also have to allocate QCs for every ship to load and unload containers on/off the AGVs. Additionally, players have to make sure the required yellow container is loaded onto the truck before a time out occurs. Consistent with real planning operations, the state of all resources has to be monitored in parallel, while making decisions about locations of QCs, RMGs, and import, export, and transfer containers. Calculations about the number of cranes needed to unload a ship in time ask for calculation skills. Many visual cues are provided on the screen to indicate the status of operations and need for actions, which ask for recognition and tracking skills. Motor skills are needed for actions such as dragging the cranes to the exact right positions. In this sense, the skills needed for container terminal planning are quite similar to the skills needed for the tasks in the MATB-II test.
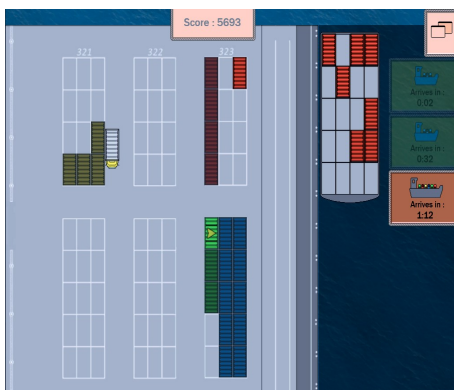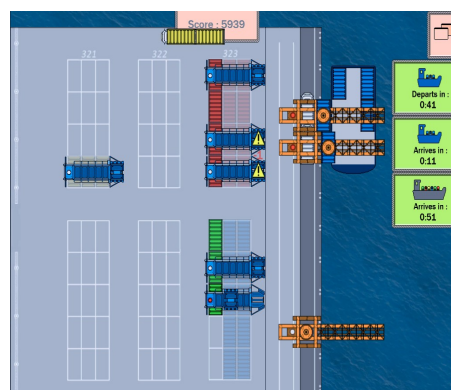
Figure 4: Planning mode of the YCS

Figure 5: Operational mode of YCS

Scoring and rules

13

All the tasks in the YCS game are interdependent. If one of the planning or resource allocation tasks is not complete, a container cannot be handled. This is also the case in reality. Coordination and alignment of all tasks is the main challenge for players. In order to stress the importance of alignment of tasks and integrated planning, and to represent the tangible effects on the performance of the container terminal, the scoring of YCS is designed as follows.

The game begins with an initial score of 1,000 points. The score of the YCS game is broken down into efficiency of handling containers, cranes (resources), and vessels (ships). An invisible score is represented by bonus points awarded for the trucks picking up import containers.

Containers. The number of containers handled within a time shift is important for the overall throughput of the container terminal. Therefore in the game, 100 points are awarded for each container that is successfully handled, meaning that the container is loaded off a ship and placed in the yard and vice-versa. For every container that was not handled a penalty of 100 points is given to the player.

Cranes (Resources). Optimal resources utilization is crucial for the performance of the container terminal. As discussed in the introductory section, one of the issues faced by container terminals is the over-utilization of some resources and under-utilization of others. To emphasize these issues, for every resource (QC, RMG, or AGV) that is idle, 1 penalty point is given for every second it remains idle. If a resource is idle, then a yellow warning sign appears on top of the QC or RMG to notify the player.

Vessels (Ships). Ship turnaround time is key for terminal business, since customer satisfaction is heavily dependent on it. Therefore, in the game, for every second the ship leaves earlier than the scheduled departure time, 100 points are awarded per container on the ship.

Bonus. Efficient and timely transfer of containers onto the hinterland is critical to maintain smooth yard operations. Therefore the trucks that arrive in the yard to pick up the containers bound towards the hinterland need to be handled within the limited time span. Each truck that leaves the yard with an import container gives the player 100 bonus points.

The overall score provides a good insight in how well participants balanced the different tasks and were able to accomplish a quick turn-around time of the ship, smooth hinterland operations, effective usage of the yard, and maximum usage of the QCs, RMGs, and AGVs. This score is used as the dependent variable in the hypotheses, and represents the success of planning operations.

Validation and testing of YCS

The YCS game has been tested and validated with experts in the port industry to establish its value as a research instrument to study planning task performance in container terminal operations. Expert validation is a recommended method to validate simulation games (Peters, Vissers & Heijne, 1998)

We sent an online validation survey to five subject matter experts to evaluate the effectiveness of YCS for training employees and new recruits in container terminal operations. The profiles of these experts are described in Table 1. These experts played the YCS game before answering the survey.

Table 1 YCS validation study: Expert profiles

Role	Years of professional experience
1. General Manager, Terminal Operating Systems	28
2. General Manager, Operations, Africa & Middle East	23
3. Managing Director - Information Technology (IT)	20
4. Technical Project Engineer (IT), Marine Terminal Operations	8
5. Consultant (personnel training) for the port industry	6

Four out of five experts agreed that the reality of container terminal operations was sufficiently represented in the game. All five experts agreed that they could relate the challenges in the game to those they encountered or observed in their real life experience. Four of the five experts stated that the processes in the container terminals were well represented given the time-frame of the game. All the five experts who evaluated the game stated that the interdependency in the processes are well represented in the game. They all agreed on the appropriateness of the scoring schema for containers, ship turn around time, and bonus points for the trucks. However all 5 experts disagreed on the negative points for idle resources and found it unrealistic. This evaluation was done on the final version of the game, so we weighed the benefits of removing the negative score and decided to keep it since the negative points were insignificant relative to the overall score.

Four out of five experts strongly agreed that the game could be used for training and evaluation of employees and new recruits involved in integrated planning operations. Expert 5 stated "I have high expectations of the game's ability to transmit the message (complexity of integrated planning operations) in a nutshell. We should be able to do a much better job than with conventional means (lectures, manuals) and it should be a great complementation to our controls trainings." Expert 1 noted "The approach of gaming (for training and recruitment) is an excellent one".

The YCS game is currently being used by a leading container terminal in the Netherlands for training purposes on their global internal training platform. Recent data from the terminal management (Feb 2017) suggests that the YCS game was played more than 15,000 times by their employees across their European, Middle Eastern and African offices. This data together with our interactions with subject matter experts support that YCS game is a reliable instrument for measuring planner task performance in container terminals.

The following section reports on the analysis of data collected during the experimental sessions with the three instruments (TIPI, MATB-II and YCS) to test the hypotheses.

6. Results

The hypotheses H1 and H2 were tested using hierarchical regression. The purpose of multiple regression is to assess the relationship between a group of independent variables on a dependent variable, after accounting the effects of other control variables on the dependent variable. In our case, our control variables are age and gender. Multi-tasking ability represented by MATB_COMP score is an independent variable, while personality type is considered to be a moderating variable affecting the relationship between multi-tasking ability and planner task performance. The dependent variable is the planner task performance as measured by the YCS score.

The results of the regression are reported in Table 2. Model 1 shows the results for age and gender, which were included as control variables, and indicates that gender has a significant influence on performance. In Model 2, the MATB score was added as a test of hypothesis 1, and in Model 3(a-e) the interaction construct relating to personality type is included as a test of hypothesis 2. As the Big Five consists of five independent traits, Model 3 investigated the interaction effects of each of the personality traits on the relationship between multi-tasking ability and planner task performance separately. The implications for the hypotheses are discussed in detail in the following sub-sections.

6.1. Multi-tasking ability and planner task performance

Revisiting the first hypothesis,

H1: Multi-tasking ability is positively related to planner task performance for container terminal operations

As we can observe from Model 2 in Table 2, 25.7% of the variance in planner task performance (represented by the YCS score) is explained by two significant variables—gender and multi-tasking ability. Gender explains 11.7% of the variance ($p<0.01$), where female participants outperform male participants. The average YCS score of female participants is 6170.92, with a standard deviation of 2968.97, whereas the average YCS score of the male participants is 4201.2 with a standard deviation of 2801.87. The age of the participants has no significant effect on the planner task performance. The participants belong to a homogenous age range, since they are students in the bachelor's and master's programs of the American and Dutch universities. An additional 14% of the variance that is statistically significant ($p<0.01$) is explained by the multi-tasking ability represented by the MATB score. Multi-tasking ability has a significant positive effect on the planner task performance for container terminals. Therefore hypothesis H1 is well supported.

6.2. Moderating effect of personality traits

The second hypothesis relating to personality traits was formulated as:

H2: Personality traits, a) Conscientiousness b) Emotional stability c) Agreeableness d) Extraversion e) Openness to experience have a moderating effect on the relationship between multitasking ability and planner task performance for container terminal operations.

Model 3 in Table 2 captures the moderating effects of personality traits. Model 3(a) examines the moderating effect of conscientiousness, and shows that this construct does not explain significant additional variance in the planner task performance. Similarly Models 3(b), 3(c), 3(d) examine the moderating effect of emotional stability, agreeableness, and extraversion, respectively; the results show these constructs do not explain any significant additional variance. Given the lack of a significant moderating effect of conscientiousness, emotional stability, agreeableness, and extraversion, there is insufficient evidence to support hypotheses H2a, H2b, H2c, H2d, therefore they are rejected.

Openness to experience as a moderating variable, shown in Model 3(e), explains an additional 2.4% of the variance. This moderating effect is statistically significant ($p<0.05$) and negative. Therefore, hypothesis 3(e), which states that openness to experience has a moderating effect on the relationship between multi-tasking ability and planner task performance is supported.

The negative moderation/interaction effect of openness to experience is illustrated in a graph as shown in Figure 6. The dependent variable planner task performance (YCS score) is on the y-axis, while the moderating variable, openness to experience is on the x-axis. Multi-tasking ability (MATB score) of the participants is divided into 3 categories:

1. High multi-tasking ability, if the MATB score is higher than 1 standard deviation than the average MATB score. It is represented by the red color.
2. Medium multi-taking ability, if the MATB score is between -1 and 1 standard deviation of the average MATB score. It is represented by the green color.
3. Low multi-taking ability, if the MATB score is below -1 standard deviation of the average MATB score. The color blue represents it.

The scatter plot of the graph in Figure 6 shows the three categories of multi-tasking ability. The graph demonstrates that in most cases, high multi-tasking corresponds to high planner task performance. We can observe from Figure 6 that majority of the participants with YCS score higher than 8000 have high multi-tasking ability.

Figure 6: Interaction effects: MATB and personality traits vs. YCS score
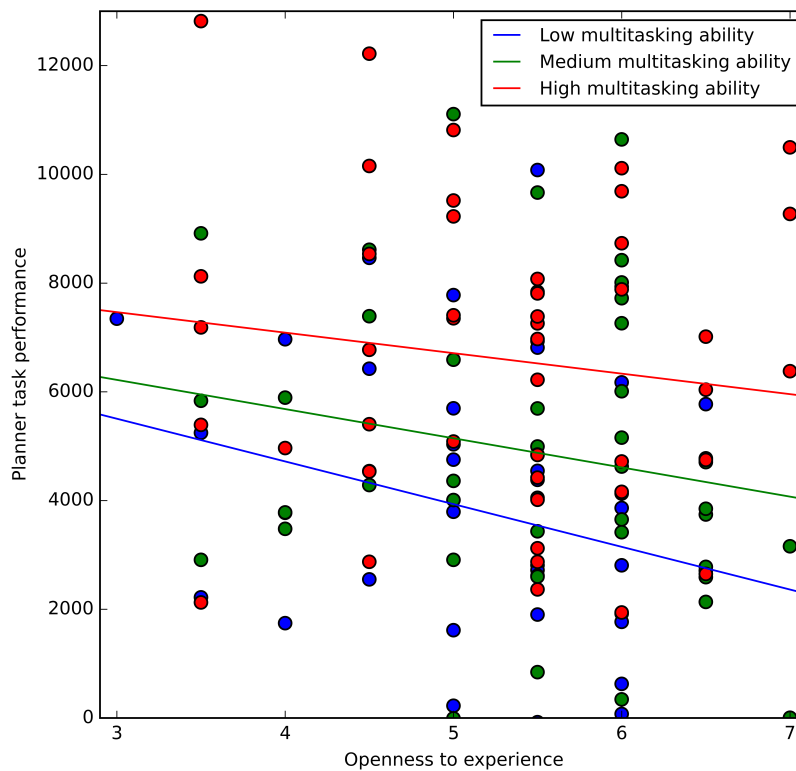
However, players with high multi-tasking ability perform slightly worse on planner task performance when they have high openness to experience. A red regression line with a negative slope shows this trend.

The same negative trend is observed for participants with medium (green line) and low (blue line) multi-tasking ability.

The results show that, for a given multi-tasking ability, the planner task performance is negatively affected by the increase in openness to experience. This negative moderating effect of openness to experience explains 2.4% of the variance in planner task performance.

We will further discuss the results and provide some conclusions and recommendations in the following section.

Table 2: Hierarchical regression: Personality trait as a moderating variable

DV: YCS score

Variables	Model 1 Controls		Model 2 Multi-tasking ability		Model 3(a) Conscientious		Model 3(b) Emotional Stability		Model 3(c) Agreeableness		Model 3(d) Extraversion		Model 3(e) Openness to Experience	
	β	t-value	B	t-value	β	t-value	β	t-value	β	t-value	β	t-value	β	t-value
Step 1														
Age	-0.109	-1.297	-0.119	-1.547	-0.116	-1.503	-0.112	-1.443	-0.124	-1.608	-0.114	-1.478	-0.114	-1.494
Gender	-0.344	-4.113**	-0.259	-3.283**	-0.265	-3.344**	-0.241	-2.969**	-0.240	-2.986**	-0.243	-3.111**	-0.260	-3.337**
Step 2														
Multi-tasking ability			0.384	4.926**	0.284	2.116*	0.290	2.433**	0.500	4.091**	0.44	3.967**	0.633	4.397**
Step 3														
Multi-tasking ability x Personality traits					0.121	0.922	0.126	1.034	-0.147	-1.231	-0.102	-0.990	-0.294	-2.049*
R^2	0.117		0.257		0.262		0.263		0.266		0.263		0.281	
Change in R^2	0.117		0.140		0.005		0.006		0.009		0.006		0.024	
F value of change in R^2	8.630**		24.265**		0.850		1.070		1.514		0.980		4.197*	

* $p < 0.05$
** $p < 0.01$

7. Discussion and Conclusions

The optimal selection, allocation and training of employees at marine container terminals plays a crucial role in achieving a high level of productivity and competitiveness of container terminals (Legato & Monaco, 2004). Planners within container terminals have a direct effect on terminal performance as they plan and manage resource use. Task performance of planners, therefore, has a direct impact on organizational goal accomplishment (Morgeson et al., 2005). Modern container terminals are driven by technology and market developments such as automation, commodity mix, and horizontal and vertical integration of terminal logistics, increasing the complexity and dynamics in logistic operations and planning tasks. Adaptive planning is highly desirable for employees to handle the complexity and dynamics of continuously changing tasks in complex systems (Rouse et al., 1992). Although the need for adaptive and integrated planning tasks of container terminal operations is well established, the skills and traits of planners that predict superior performance at such tasks are largely unknown. Our research attempted to close this gap. We analyzed the relationship between multi-tasking ability and personality type on planner task performance.

Based on our results we can conclude that multi-tasking ability has a strong relation with superior planner task performance. This finding is consistent with the crucial role of multi-tasking ability on performance in other domains. Irrespective of the domain, operators in complex systems need to coordinate the performance of multiple tasks (Chiappe, Vu, & Strybel, 2012; Liu, Wasson, & Vincenzi, 2009). Cognitive demands on operators are likely to increase in the future (Chiappe et al., 2012; Liu et al., 2009). Although automation reduces physical workload, it increases cognitive workload (McCarley & Wickens, 2005; Dehais, Causse, Vachon, & Tremblay, 2012). Fortunately, recent studies have found that capacity for attention to perform multiple tasks can be enhanced (Bavelier, Achtman, Mani, & Föcker, 2012). Some of the suggested training instruments include simulation games (Bavelier et al., 2012, Chiappe et.al, 2013). Therefore, a key recommendation from our study to employers like container terminal operators is to provide training programs to employees to enhance their multi-tasking ability to perform well under a complex and dynamic workload.

Additionally, we found a significant gender difference in planner task performance, measured by the YCS game. We approach an interpretation of this difference from two perspectives. First, this difference could be explained by different performance of male and female users in digital simulation games, or virtual environments in general. Unfortunately, related research cannot provide an explanation for this, as most studies show that in general, male individuals perform better in virtual environments and games and use them more often. For example, differences in spatial abilities in VR are shown in the studies of Astur, Ortiz, and Sutherland (1990), Sandstrom, Kauffman, and Huettel (1998), and Cutmore, Hine, Maberly, Langford, and Hawgood (2000). In these studies, male test persons perform better in spatial orientation than female participants. One reason for this is found in the level of anxiety that test persons show when learning to navigate through a virtual environment, with a higher level of anxiety in the female population (Bryant, 1982; Lawton, 1994; Lawton, 1996). Results from another study also indicate that gender differences in favor for the male users also emerge from a different ability in the use of interfaces (Waller, 2000). Amongst others, Imhof, Vollmeyer, and Beierlein (2007) showed a higher use of computers, and in particular, internet applications, by male students across Europe. They also show that male students use the computer much more often to play games than female students. The same study shows that gender has a direct impact on computer performance, with the male users outperforming the females in a task including PowerPoint, which cannot be explained by the authors themselves. A collection of game-related studies shows that, while the number of female players of digital games has increased in recent years, the content of games still seems to address male players (Kafai, Heeter, Denner, & Sun, 2008). The gender differences in socialization towards the use of computers is also shown by Whitley (1996) and could be one reason for actual differences found in the studies mentioned above. In summary, the focus on

computer use and simulation game performance does not provide us with an explanation for the gender differences we found in our study. In general, related work suggests the opposite, namely male test persons performing better in computer use, virtual environments and games, based on lower anxiety level and a more positive socialization towards the use of digital media.

Secondly, we were looking for an explanation for the gender difference in planning performance from the perspective of planning performance itself. In their study, Naglieri and Rojahn (2001) find a better performance of girls in planning and attention scales as result of a test based on the Planning, Attention, Simultaneous, Successive (PASS) cognitive-processing theory. In more general terms than focused on virtual environments only, cognitive research seems to be consistent in studies on gender differences related to specific patterns of intellectual abilities. Women in these studies perform better on verbal and memory tasks, while men do better on spatial cognition and spatial learning (see for an overview Boghi et al., 2006). The same study shows that men and women use different strategies in a planning task, which was represented in their brain activity. It cannot be said whether this is a biological or a social phenomenon, with both genders learning different approaches towards complex planning tasks throughout their life (Boghi et al., 2006). From these studies, we can draw the conclusion that different cognitive approaches towards planning tasks could be an explanation for the differences found in our study, too. Still, there is plenty of variance unexplained between men and women that might rather be related to individual differences than to gender differences (Unterrainer, Ruff, Rahm, Kaller, Spreer, Schwarzwald, & Halsband, 2005).

In summary, we can say that planning ability could be one explanation for the gender difference we found in our study. This is at least more probable than the influence of the test instrument being a computer game. Nonetheless, being able to provide a clear statement on the gender differences in general would mean to conduct further analyses thoroughly grounded on related theories, to include as well neurological as sociocultural aspects that could influence the differences found (Naglieri & Rojahn, 2001). We are aware of the fact that being a woman or a man does not necessarily mean that the one or other performs better or worse based only because of her or his biological sex, but that social and individual differences also play a very important or even more important role (Unterrainer et al., 2005). The measurement of personality traits included in our study investigates an interaction effect along this assumption.

The second part of our results demonstrates that the multi-tasking skills overpower the personality traits with one notable exception. Indeed, each of the personality traits was tested as an interaction effect with the multi-tasking variable to determine whether each trait intensified or lessened the impact of multi-tasking on performance. The only trait with a significant impact was the openness trait. Openness as a trait significantly lessens the positive impact of multi-tasking. Previous studies have positively associated openness to experience with superior performance of creative tasks (Baer & Oldham, 2006; LePine, Colquitt, & Erez, 2000; George & Zhou, 2001). The planning tasks in container terminals are not purely creative in nature, however, as they are bound by rules and deadlines. Although we cannot make a strong recommendation regarding this personality trait, employers may benefit from considering personality traits as well as other skills such as multi-tasking ability while allocating tasks to employees, especially new recruits. Other studies have already underpinned the use of the Big Five model for employee selection generally (see e.g. the meta-review of Hurtz & Donovan, 2000), and more specifically for logistics personnel involved in marketing tasks (Periatt, Chakrabarty, & Lemay, 2007). With a combination of tools like we used in our study, and the integration of skills like multi-tasking ability, the use of the model for employee selection and training in logistics planning activities could be further enhanced.

In addition to the results from our experimental session, this study also highlights the potential of simulation games as a research method for empirical studies in transportation.

The most common research methods in the field of supply chain, transportation, and logistics are case-studies, surveys, action-research, modelling, and simulation (Meijer, 2009; Bryman, 2015; Voss, Tsikriktsis & Frohlich, 2002). However simulation gaming has a strong potential to join the above list (Bryman, 2015). The key disadvantages of surveys are the level of truthfulness of responses and involvement of participants. Questions may be interpreted differently by different respondents. Case studies have the issue of low repeatability, while action research is time consuming and also struggles with lack of control over the research environment. Case-studies, surveys and action research can be categorized as analytical approaches (Bryman, 2015), while simulation gaming, modelling, and simulation can be placed both in the design sciences and in the analytical sciences (Kriz & Hense, 2006). From the dual stance, analytical science uses computer simulations and simulation gaming as research methods to test theories in various fields, whereas the design science emphasizes on the design of the artefact, and testing its usability (Klabbers, 2009; Kriz & Hense, 2006). In our case, we designed the simulation game for the purpose of studying planner task performance in container terminal operations. We tested and validated the game and used it to test two hypotheses related to the role of multi-tasking ability and personality type on planner task performance.

Modelling and simulation offer unlimited testing possibilities but lack the tacit input related to human behavior. Simulation games, although less realistic than simulation models, have the ability to engage human participants in a safe and controlled experimental setting and record their behavior and decision making. This has been demonstrated in our research approach.

In addition to its research value, the YCS game has already been adopted as a professional training tool by a leading container terminal. This also reflects the interest and enthusiasm of the port industry to use games for evaluation and training of employees.

Although the research was conducted in a rigorous manner, it is not without limitation. The results showed that four of the personality traits (conscientiousness, emotional stability, agreeableness, and extraversion) did not have a moderating impact on the relationship between multi-tasking skills and performance. It should be noted, however, as a limitation that the YCS game is a single player game, while skills such as agreeableness and extraversion may only be observable in a multi-player setting. Thus, as we incorporate a multi-player component in our simulation environment, we will be providing a more realistic assessment of the moderating impact of these personality traits on multi-tasking skills. The sample population for the experiments were students although their specialization was in the field of logistics, supply chains and transportation. In addition, the level of complexity of the YCS game is not as high as complexity in a real conainer terminal. Still, the results clearly indicate the relation between multi-tasking and planner performance. We expect this to become even more important when more complex and realistic planning tasks closer to the operations in real container terminals are studied. We plan to research this in follow-up studies using high-fidelity simulation models of container terminals and emulated planning and control systems.

8. Acknowledgment

References

Adler, R. F. & Benbunan-Fich, R. (2012). Juggling on a high wire: Multitasking effects on performance. *International Journal of Human-Computer Studies, 70*(2), 156–168.

Appelbaum, S. H., Marchionni, A. & Fernandez, A. (2008), The multitasking paradox: Perceptions, problems and strategies. *Management Decision, 46*(9), 1313–1325.

Astur, R. S., Ortiz, M. L., & Sutherland, R. J. (1998). A characterization of performance by men and women in a virtual Morris water task: A large and reliable sex difference. *Behavioural Brain Research*, *93*(1), 185–190.

Baer, M., & Oldham, G. R. (2006). The curvilinear relation between experienced creative time pressure and creativity: moderating effects of openness to experience and support for creativity. *Journal of Applied Psychology, 91*(4), 963–970.

Baines, T. S., Asch, R., Hadfield, L., Mason, J., Fletcher, S. & Kay, J. M. (2005). Towards a theoretical framework for human performance modelling within manufacturing systems design. *Simulation Modelling Practice and Theory, 13*(6), 486–504.

Baron, R. M., & Kenny, D. A. (1986). The moderator–mediator variable distinction in social psychological research: Conceptual, strategic, and statistical considerations. *Journal of Personality and Social Psychology*, 51(6), 1173–1182.

Bavelier, D., Achtman, R. L., Mani, M., & Föcker, J. (2012). Neural bases of selective attention in action video game players. *Vision Research, 61*, 132–143.

Boghi, A., Rasetti, R., Avidano, F., Manzone, C., Orsi, L., D'agata, F., Caroppo, P., Bergui, M., Rocca, P., Pulvirenti, L., Bradac, G. B., Bogetto, F. Mutani, R. & Mortara, P. (2006). The effect of gender on planning: An fMRI study using the Tower of London task. *Neuroimage*, *33*(3), 999–1010.

Bryant, K. J. (1982). Personality correlates of sense of direction and geographic orientation. *Journal of Personality and Social Psychology*, *43*(6), 1318–1324.

Bryman, A. (2015). *Social research methods*. Oxford, UK: Oxford University Press.

Bühner, M., König, C. J., Pick, M., & Krumm, S. (2006). Working memory dimensions as differential predictors of the speed and error aspect of multitasking performance. *Human Performance, 19*(3), 253–275.

Burke, C. S., Pierce, L. G., & Salas, E. (Eds.). (2006). *Understanding adaptability: A prerequisite for effective performance within complex environments*. Amsterdam: Elsevier.

Caldwell, J. A., & Ramspott, S. (1998). Effects of task duration on sensitivity to sleep deprivation using the multi-attribute task battery. *Behavior Research Methods, Instruments, & Computers, 30*(4), 651–660.

Carlo, H. J., Vis I. F. A., & Roodbergen, K. J. (2014). Storage yard operations in container terminals: Literature overview, trends, and research directions. *European Journal of Operational Research, 235*, 412–430.

Chiappe, D., Vu, K. P. L., & Strybel, T. (2012). Situation awareness in the NextGen air traffic management system. *International Journal of Human-Computer Interaction, 28*(2), 140–151.

Comstock, J. R., & Arnegard, R. J. (1992). *The Multi-Attribute Task Battery for human operator workload and strategic behavior research*. Hampton, VA: National Aeronautics and Space Administration, Langley Research Center.

Cutmore, T. R., Hine, T. J., Maberly, K. J., Langford, N. M., & Hawgood, G. (2000). Cognitive and gender factors influencing navigation in a virtual environment. *International Journal of Human-Computer Studies, 53*(2), 223–249.

Dehais, F., Causse, M., Vachon, F., & Tremblay, S. (2012). Cognitive conflict in human–automation interactions: a psychophysiological study. *Applied Ergonomics*, *43*(3), 588–595.

De Oliveira, G. F., & Cariou, P. (2015). The impact of competition on container port (in) efficiency. *Transportation Research Part A: Policy and Practice, 78*, 124-133.

Digman, J. M. (1989). Five robust trait dimensions: Development, stability, and utility. *Journal of Personality, 57*(2), 195–214.

Digman, J. M. (1990). Personality structure: Emergence of the five-factor model. *Annual Review of Psychology, 41*, 417–440.

Fairclough, S. H., Venables, L., & Tattersall, A. (2005). The influence of task demand and learning on the psychophysiological response. *International Journal of Psychophysiology*, 56(2), 171–184.

Foehr, U. G. (2006). *Media multitasking among American youth: Prevalence, predictors and pairings*. Menlo Park, CA: Henry J. Kaiser Family Foundation.

Gambardella, L. M., Rizzoli, A. E. & Zaffalon, M. (1998). Simulation and planning of an intermodal container terminal. *Simulation, 71*(2), 107–116.

George, J. M. & Zhou, J. (2001). When openness to experience and conscientiousness are related to creative behavior: an interactional approach. *Journal of Applied Psychology*, *86*(3), 513–524.

Goldberg, L. R. (1990). An alternative "description of personality": the big-five factor structure. *Journal of Personality and Social Psychology, 59*(6), 1216–1229.

Gordon, J. R., Lee, P.M. & Lucas, H. C. (2005). A resource-based view of competitive advantage at the port of Singapore. *The Journal of Strategic Information Systems 14*(1), 69–86.

Gosling, S. D., Rentfrow, P. J. & Swann, W. B. (2003). A very brief measure of the big-five personality domains. *Journal of Research in Personality, 37*(6), 504–528.

Harteveld, C. (2011). *Triadic Game Design: Balancing Reality, Meaning and Play*. London: Springer Science & Business Media.

Harteveld, C. & Sutherland, S. C. (2015). The goal of scoring: exploring the role of game performance in educational games. *In Proceedings of the 33rd Annual ACM Conference on Human Factors in Computing Systems* (pp. 2235–2244). New York: ACM.

Harris, W. C., Hancock, P. A., Arthur, E. J., & Caird, J. K. (1995). Performance, workload, and fatigue changes associated with automation. *The International Journal of Aviation Psychology, 5*(2), 169–185.

Hurtz, G. M. & Donovan, J. J. (2000). Personality and job performance: the big five revisited. *Journal of Applied Psychology, 85*(6), 869–879.

Imhof, M., Vollmeyer, R., & Beierlein, C. (2007). Computer use and the gender gap: The issue of access, use, motivation, and performance. *Computers in Human Behavior*, *23*(6), 2823–2837.

Kafai, Y. B., Heeter, C., Denner, J., & Sun, J. Y. (2008). *Beyond Barbie [R] and Mortal Kombat: New Perspectives on Gender and Gaming*. Cambridge (MA), MIT Press (BK).

Klabbers, J. H. (2009). *The magic circle: Principles of gaming & simulation*. Rotterdam, NL: Sense Publishers.

Kriz, W. C., & Hense, J. U. (2006). Theory-oriented evaluation for the design of and research in gaming and simulation. *Simulation & Gaming*, 37(2), 268-283.

Kurapati, S., Groen, D., Lukosch, H. & Verbraeck, A. (2014). Microgames in practice: A case study in container terminal operations. In *Kriz, W. (Ed): The Shift from Teaching to Learning: Individual, Collective and Organizational Learning Through Gaming Simulation* (pp. 333–346). Dornbirn, Austria: ISAGA.

Lawton, C. A. (1994). Gender differences in way-finding strategies: Relationship to spatial ability and spatial anxiety. *Sex Roles*, *30*(11-12), 765–779.

Lawton, C. A. (1996). Strategies for indoor wayfinding: The role of orientation. *Journal of Environmental Psychology*, *16*(2), 137–145.

Legato, P. & Mazza, R. M. (2001). Berth planning and resources optimisation at a container terminal via discrete event simulation. *European Journal of Operational Research, 133*(3), 537–547.

Legato, P. & Monaco, M. F. (2004). Human resources management at a marine container terminal. *European Journal of Operational Research, 156*(3), 769–781.

LePine, J. A., & Van Dyne, L. (2001). Voice and cooperative behavior as contrasting forms of contextual performance: evidence of differential relationships with big five personality characteristics and cognitive ability. *Journal of Applied Psychology*, *86*(2), 326–336.

LePine, J. A., Colquitt, J. A., & Erez, A. (2000). Adaptability to changing task contexts: Effects of general cognitive ability, conscientiousness, and openness to experience. *Personnel Psychology, 53*(3), 563–593.

Liu, D., Wasson, R., & Vincenzi, D. A. (2009). Effects of system automation management strategies and multi-mission operator-to-vehicle ratio on operator performance in UAV systems. *Journal of Intelligent and Robotic Systems, 54*(5), 795–810.

Lu, C. & Kuo, S. (2016). The effect of job stress on self-reported safety behavior in container terminal operations: The moderating role of emotional intelligence. *Transportation Research: Part F*, 37, 10-26.

McAfee, B., Glassman, M., & Honeycutt, E. D. (2002). The effects of culture and human resource management policies on supply chain management strategy. *Journal of Business Logistics*, 23(1), 1-18.

McCarley, J. S., & Wickens, C. D. (2005). *Human factors implications of UAVs in the National Airspace*. University of Illinois at Urbana-Champaign: Aviation Human Factors Division.

Meijer, S. (2009). *The organisation of transactions. Studying supply networks using gaming simulation.* PhD dissertation. Wageningen University, The Netherlands.

Meisel, F. (2009). *Seaside operations planning in container terminals*. Berlin, Heidelberg: Physica-Verlag.

Morgeson, F. P., Reider, M. H., & Campion, M. A. (2005). Selecting individuals in team settings: The importance of social skills, personality characteristics, and teamwork knowledge. *Personnel Psychology*, *58*(3), 583–611.

Murty, K. G., Liu, J., Wan, Y., &, Linn, R. (2005) A decision support system for operations in a container terminal. *Decision Support Systems, 39*(3), 309–332.

Niehaus, J., & Riedl, M. O. (2009). Scenario adaptation: An approach to customizing computer-based training games and simulations. In *Proceedings of the AIED 2009 Workshop on Intelligent Educational Games* (Vol.3, pp. 89–98). Brighton: AIED.

Notteboom, T. (2012). Dock labour systems in North-West European seaports: how to meet stringent market requirements?. In *Proceedings of the International Forum on Shipping, Ports and Airports* (pp. 365– 383). Hong Kong: IFSPA.

Oberauer, K. and Kliegl, R. (2004). Simultaneous cognitive operations in working memory after dual-task practice. *Journal of Experimental Psychology*, 30(4), 689–707.

Ophir, E., Nass, C., & Wagner, A. D. (2009). Cognitive control in media multitaskers. *Proceedings of the National Academy of Sciences*, 106(37), 15583–15587.

Othman, M., Bhuiyan, N. & Gouw, G. J. (2012). Integrating workers' differences into workforce planning. *Computers & Industrial Engineering, 63*(4), 1096–1106.

Pashler, H. (2000). Task switching and multitask performance. In: Monsell, S. & Driver, J. (Eds.): *Attention and Performance XVIII: control of mental processes.* Cambridge, MA: MIT Press, 275–307.

Penney, L. M., David, E. & Witt, L. (2011). A review of personality and performance: Identifying boundaries, contingencies, and future research directions. *Human Resource Management Review*, *21*(4), 297–310.

Periatt, J. A., Chakrabarty, S., & Lemay, S. A. (2007). Using personality traits to select customer-oriented logistics personnel. *Transportation Journal*, 46(1), 22-37.

Pervin, L. A., & John, O. P. (1999). *Handbook of personality: Theory and research*. New York: The Guildford Press.

Peters, V., Vissers, G. & Heijne, G. (1998). The validity of games. *Simulation & Gaming*, *29*(1), 20–30.

Pew, R. W. & Mavor, A. S. (Eds.) (1998). *Modeling human and organizational behavior: application to military simulations*. Washington, D.C: National Academy Press.

Pulakos, E. D., Arad, S., Donovan, M. A. & Plamondon, K. E. (2000). Adaptability in the workplace: development of taxonomy of adaptive performance. *Journal of Applied Psychology, 85*(4), 612–624.

Raser, J. R. (1969). *Simulation and society: an exploration of scientific gaming*. Boston: Allyn & Bacon.

Rose, E. (2010). Continuous partial attention: Reconsidering the Role of online learning in the age of interruption. *Educational Technology Magazine: The Magazine for Managers of Change in Education*, *50*(4), 41–46.

Rotundo, M. & Sackett, P. R. (2002). The relative importance of task, citizenship, and counterproductive performance to global ratings of job performance: a policy-capturing approach. *Journal of Applied Psychology, 87*(1), 66–80.

Rouse, W. B., Cannon-Bowers, J., & Salas, E. (1992). The role of mental models in team performance in complex systems. *IEEE Transactions on Systems, Man and Cybernetics, 22*(6), 1296–1308.

Salvucci, D. D. & Taatgen, N. A. (2010). *The multitasking mind*. New York: Oxford University Press.

Sandstrom, N. J., Kaufmann, J., & Huettel, S. A. (1998). Males and females use different distal cues in a virtual environment navigation task. *Cognitive Brain Research*, 6, 351–360

Santiago-Espada, Y., Myer, R., Latorella, K. & Comstock, J. (2011). *The Multi-Attribute Task Battery II (MATB-II) software for human performance and workload research: A user's guide* (NASA/TM-2011–217164). Hampton, Virginia: NASA Langley Research Center.Stone, L. (2007). Continuous Partial Attention. Retrieved on 11.01.2016 from: http://lindastone.net/qa/continuous-partial-attention/

Swart, W., Hall, C., & Chen, H. (2012). Human performance in supply chain management. *Supply Chain Forum: An International Journal*, 13(2), 10-20.

Thomas, R. (2011). When student samples make sense in logistics research. *Journal of Business Logistics*, 32(3), 287-290.

Turnbull, P. & Weston, S. (1993). The British port transport industry part 2. Employment, working practices and productivity. *Maritime Policy and Management, 20*(3), 181–195.

UNCTAD (2014). *Review of maritime transport*. Technical Report UNCTAD/RMT/2014. Geneva: United Nations.

Unterrainer, J. M., Ruff, C. C., Rahm, B., Kaller, C. P., Spreer, J., Schwarzwald, R., & Halsband, U. (2005). The influence of sex differences and individual task performance on brain activation during planning. *Neuroimage*, *24*(2), 586–590.

Voss, C., Tsikriktsis, N., & Frohlich, M. (2002). Case research in operations management. *International journal of operations & production management*, 22(2), 195-219.

Waller, D. (2000). Individual differences in spatial learning from computer-simulated environments. *Journal of Experimental Psychology: Applied*, *6*(4), 307–321.

Whitley, B.E. (1996). Gender differences in computer-related attitudes and behavior: a meta-analysis. *Computers in Human Behavior*, 13(1), 1–22

Yoon, J., Lee, H. Y., & Dinwoodie, J. (2015). Competitiveness of container terminal operating companies in South Korea and the industry–university–government network. *Transportation Research Part A: Policy and Practice, 80*, 1-14.